



A Model-Based Approach for Compound Leaves Understanding and Identification

Guillaume Cerutti, Laure Tougne, Julien Mille, Antoine Vacavant, Didier
Coquin

► To cite this version:

Guillaume Cerutti, Laure Tougne, Julien Mille, Antoine Vacavant, Didier Coquin. A Model-Based Approach for Compound Leaves Understanding and Identification. International Conference on Image Processing (ICIP), Sep 2013, Melbourne, Australia. pp.1471-1475. hal-00872889

HAL Id: hal-00872889

<https://hal.science/hal-00872889>

Submitted on 14 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A MODEL-BASED APPROACH FOR COMPOUND LEAVES UNDERSTANDING AND IDENTIFICATION

Guillaume Cerutti^{1,2} Laure Tougne^{1,2} Julien Mille^{1,3} Antoine Vacavant⁴ Didier Coquin⁵

¹ Université de Lyon, CNRS

² Université Lyon 2, LIRIS, UMR5205, F-69676, France

³ Université Lyon 1, LIRIS, UMR5205, F-69622, France

⁴ Clermont Université, Université d'Auvergne, ISIT, F-63001, Clermont-Ferrand

⁵ LISTIC, Domaine Universitaire, F-74944, Annecy le Vieux

ABSTRACT

In this paper, we propose a specific method for the identification of compound-leaved tree species, with the aim of integrating it in an educational smartphone application. Our work is based on dedicated shape models for compound leaves, designed to estimate the number and shape of leaflets. A deformable template approach is used to fit these models and produce a high-level interpretation of the image content. The resulting models are later used for the segmentation of leaves in both plain and natural background images, by the use of multiple region-based active contours. Combined with other botany-inspired descriptors accounting for the morphological properties of the leaves, we propose a classification method that makes a semantic interpretation possible. Results are presented over a set of more than 1000 images from 17 European tree species, and an integration in the existing mobile application Folia¹ is considered.

Index Terms— plant recognition, compound leaf, deformable templates, image segmentation, active contours, classification

1. INTRODUCTION

When considering a tree identification mobile application, leaves are an obvious choice for recognition. They can be found all year long, are plane enough to be easily photographed, and show properties that make the identification possible. However, leaves are natural objects whose morphological diversity makes modelling complicated, not to say impossible, and distinguishing between leaf types is necessary. This work concerns only compound leaves, for which a single leaf is actually divided in many leaflets.

The proposed method aims at modelling explicitly the disposition, the global shape and the local specificities of the leaflets in order to classify photographs of leaves in a natural environment into a list of species. The use of high-level, independent descriptors offers the possibility of an explanatory process, putting semantic concepts on what is recognized, which may be of great interest for the user. Figure 1 gives an insight of the proposed process.

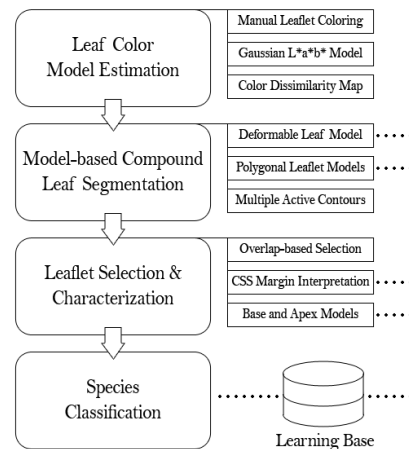


Fig. 1. Overview of the compound leaf recognition process

Section 2 presents other works addressing close topics. The specific models used to represent and segment compound leaves are introduced in Section 3. Section 4 expounds the descriptors and classification we use, as well as some results, and perspectives of future work are given in Section 5.

2. RELATED WORK

Plant recognition is a topic of interest in image processing, mostly in the context of leaf image retrieval. Some authors [1] even share our goal of conceiving a mobile application with great success², though being designed essentially for white-background images. The problem of segmenting leaves from a natural environment appears indeed to be much more challenging and is addressed by few other authors, using very complex methods [2, 3].

Segmentation of natural objects is a context where the introduction of prior shape knowledge could be very beneficial. Deformable models, which can be traced back to active contours [4], are a popular way to achieve this. Deformable templates modeling complex objects [5], strongly constrained

Research with the reference ANR-10-CORD-005 (REVES project).

¹<https://itunes.apple.com/app/fofia/id547650203>

²<http://leafsnap.com>: developed by researchers from Columbia University, the University of Maryland, and the Smithsonian Institution

templates, such as active shape models [6] or level-set active contours with shape priors [7] constitute ways of including prior knowledge in the segmentation. However, their shapes generally lack the necessary flexibility and expressiveness to capture the diversity of leaf shapes.

Concerning leaf shape description, many works apply established shape descriptors such as the Inner-Distance Shape Context [1], moments [3], Centroid-Contour Distance curves [2] or Curvature-Scale Space transform [8]. Such descriptors were not designed to take into account the nature of the object, even if they fit quite well with its specificities. On the other hand, some explicit geometric descriptions of the leaf morphology have been proposed [9, 10].

3. DEFORMABLE COMPOUND LEAF MODELS

Similarly to what we have done in the case of simple leaves [11], the segmentation method we propose relies on the prior evaluation of a flexible leaf model, designed to cover the variety of leaf shapes. Such a model constitutes a way of providing a first segmentation as well as a description of the leaf's global shape that can later be used for recognition.

3.1. Deformable Compound Leaf Model

A first model that tries to estimate the number and disposition of the leaflets was introduced in [12] and was modified here to achieve better robustness. It is crucial to point out that very often, the number of leaflets is not the number of connected components one would find in the segmented image, given that the overlap between leaflets is a constant risk. Estimating the actual number location of those leaflets beforehand would therefore be a guarantee that the resulting shapes will be correctly described.

This model makes assumptions about the axial symmetry of the leaves, and the regularity of the leaflets in size and orientation, that are not strictly speaking always true, but are satisfying for a first approximation. As shown in Figure 2, the model represents leaflets by a variable number n_L of pairs of circles $(C_{2l}, C_{2l+1})_{l=1..n_L}$ symmetrically positioned on either side of a curved axis defined by two points T (for the top leaflet C_1) and B (for the base of the petiole) and a curvature parameter k . The additional parameters used to build the model are the following :

- $(p_l)_{l=1..n_L}$, the position of pairs of circles on the axis
- d , the distance of all the circles to the axis
- r , the radius of all the circles

The estimation of the optimal model M^* on the actual image is performed through the minimization of an energy function by successive variations of the parameters. This energy term is based on a color dissimilarity map estimated beforehand (Figure 3(b)) that accounts for each pixel's likelihood of being part of the leaf, given only its color. It is

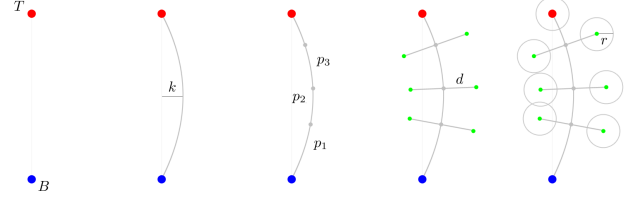


Fig. 2. Construction of the compound leaf model

based on a simple leaf color model computed after a rough coloring of the three top leaflets (Figure 3(a)) by the user. This intuitive phase corresponds to what will be asked to an user of the mobile application, and turns out to be very handy to place the model accurately in its initialization.

In the context of compound leaves, we model the color by a single Gaussian (μ, Σ) in the $L^*a^*b^*$ colorspace, which is a perceptually more accurate representation than RGB. The dissimilarity of a pixel p to this color model is then simply given by the Mahalanobis distance $d(p, \mu, \Sigma)$ with respect to the computed Gaussian parameters.

The energy function the model minimizes during its evolution is simply the sum of the dissimilarity over the region defined by the model, minus a maximal dissimilarity that acts as a balloon force:

$$E(M) = \sum_{p \in \bigcup_{i=1}^{2n_L+1} C_i} d(p, \mu, \Sigma) - d_{max} \quad (1)$$

The number of leaflets n_L has to be optimized separately, since the changes it produces in the shape of the model are too important to consider it in a gradient-descent like approach. To overcome this difficulty, the model is initialized with an excessive number of leaflets. Following a process close to simulated annealing, a temperature variable is slowly decreasing through the evolution, and brutally raised in regular cycles. At the end of each of those cycles, circles that have grouped in actual leaflets are simply suppressed, a decision made comparing the distance between the centers of two consecutive pairs of circles and the radius r . This way, unlike what was done in [12] where the number of leaflets was approximated *a posteriori*, the convergence of the model ideally shows one single circle per leaflet (Figure 3(c)).

3.2. Deformable Joint Polygonal Leaflet Models

To capture the global shape of the leaflets, we rely on the polygonal leaf model introduced in [13], making the assumption that all the leaflets share the same shape. This is generally true (and species are described in the literature by the shape of their leaflets anyway) even if little exceptions (appreciably different shapes for the top leaflets) might cause some problems.

Consequently we propose a novel joint approach where we place one model Π_i for each of the $2n_L + 1$ leaflets obtained after the evaluation of the leaf model, and constrain

them to have the same shape parameters. Only the points defining the base and apex of each leaflet vary independently. This new model $\{\Pi_i\}_{i=1}^{2n_L+1}$ evolves the same way as the previous one, minimizing the same energy function, but with no suppression of overlapping leaflets. Constraints are added throughout the evolution, on the shape parameters so that leaflets keep leaf-like shapes, and on the points so that pairs of leaflets remain locally symmetrical, under the form of an internal energy term.

The process of fitting a model to a single leaflet, which may overlap with its neighbours is risky, but the fact that all the leaflet models are evaluated simultaneously ensures that they self-constrain (Figure 3(d)). The parameters of the optimal models $\{\Pi_i^*\}_{i=1}^{2n_L+1}$ we obtain are therefore more robust than if they were computed independently on each leaflet.

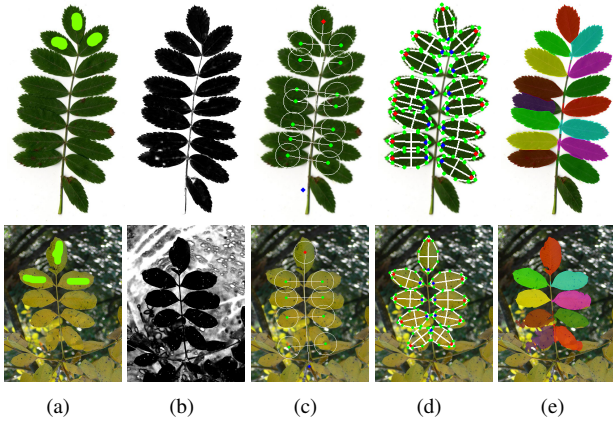


Fig. 3. Example of model-fitting and segmentation of compound leaves

3.3. Multiple Active Contour Segmentation

To obtain different interpretable contours corresponding to each leaflet, the natural choice is to deform the polygonal shapes resulting from the previous step towards the actual contours. Once again, it is very interesting to perform this step in a joint fashion, so that the contours we are evaluating act as a constraint on each other.

The contour of each leaflet is represented by a region-based active contour model, using an extension of the level-set framework to the case of multiple regions [14] and an implicit definition approximating the original level-set evolution [15]. This model have the limitation that a pixel can only be part of one region, so that the contours do not interpenetrate.

The $2n_L + 1$ regions $\{\Omega_i\}_{i=1}^{2n_L+1}$ evolve simultaneously, minimizing the energy functional :

$$E(\{\Omega_i\}_{i=1}^{2n_L+1}) = \sum_{i=1}^{2n_L+1} \omega_L E_{Leaf}(\Omega_i) + \omega_{\Pi} E_{Shape}(\Omega_i, \Pi_i^*) + \omega_{\nabla} E_{Gradient}(\Omega_i) + \omega_S E_{Smooth}(\Omega_i) - \omega_{Balloon} \quad (2)$$

The energy is composed of an external term based on the same color dissimilarity as in 3.1 and 3.2 and on the gra-

dient, and an internal term containing a shape constraint to remain close to the corresponding polygonal model Π_i^* and a smoothness term. After evolution, the result is a set of independent contours (Figure 3(e)) that can be studied to extract local properties.

4. SPECIES CLASSIFICATION & RESULTS

4.1. Shape Descriptors

The descriptors we use to learn and classify compound leaves are directly inspired from botany and aim at capturing in a decorrelated way the various morphological specificities of the leaf. This decision falls within the choice of an explanatory recognition process, where the justifications leading to the classification should be displayed to the user of the application using high-level semantic concepts. The introduced descriptors consist of:

- Parameters of the compound leaf model (3.1)
- Parameters of the polygonal leaflet model (3.2)
- Averaged parameters of base and apex models ([11])
- Averaged CSS-based contour parameters ([12])

The features describing the base, apex and margin shapes, are extracted and averaged over a subset of the leaflet contours. The idea is to discard the overlapping leaflets, for which the contour is unlikely to be accurate. To perform this selection, we compute an overlap score for every pair of leaflets, and only those which have a low score with all the others are kept. This step ensures that the extracted parameters will represent actual leaflet shapes and not erratic shapes resulting from an uncertain border between two overlapping leaflets.

4.2. Learning and Classification

We tested our methods on a subset of the PI@ntLeaves II Dataset [16] considering only species with compound leaves. The resulting set consists of 1040 images spread over 17 different species, the images being of three types : scans, scan-like (photographs on a light plain background) and photographs.

The database formed by all the parameters for all the examples is first normalized and centered. The classification algorithm is rather naive, as it simply computes for a given species \mathcal{S} a class model $\Phi_{\mathcal{S}}$ per species, keeping only the mean $\mu_{\mathcal{S},p}$ and standard deviation $\sigma_{\mathcal{S},p}$ of each parameter over all the examples for the species.

To classify a new example, we compute the distance of its parameter vectors P to all the class models. Rather than using the Euclidean distance that does not account at all for intra-class variability, or the Mahalanobis distance that distorts the parameter space and penalizes classes with low variability, we estimate the distance to the surface of the ellipsoid defined by the means and standard deviations in the parameter space.

This distance is given by:

$$\mathcal{D}(P, \Phi_S) = \|P - \mu_S\|_2 \max \left(1 - \frac{1}{\|P - \mu_S\|_M}, 0 \right) \quad (3)$$

In this equation $\|P - \mu_{S,n_L}\|_M$ is the Mahalanobis distance, which is simply in our diagonal case a normalized Euclidean distance.

These distances are computed separately for the 4 sets of descriptors presented in 4.1, which are meant to represent independent leaf features. Each of those distance terms is then weighted accordingly to its significance, estimated by observing the average distance to the correct class over the training base. The sum of the weighted distances is used to produce a ranked list of species, the top five of which are presented to the user as a result.

4.3. Results & Evaluation

We measured our results in terms of correct classification rate, not only for the first species of the resulting list, but also for the top k answers, k ranging from 1 to 5. This is a way to evaluate the performance relatively to the application of our method in a pertinent way. The classification rates are 61% on scan images, 60% on scan-like images and 43% on photographs. One interesting fact is that the presence of the correct species in the list presented to user climbs up to 95% in the case of plain background images, and 86% for natural environment photographs. A more detailed view of these results can be seen in Figure 4.

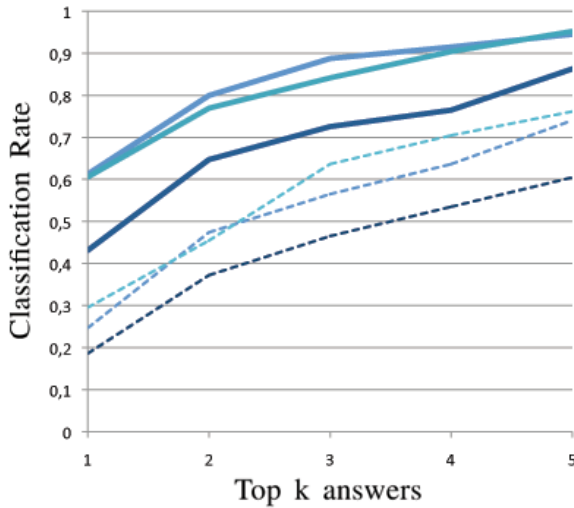


Fig. 4. Classification scores for scan (—), pseudoscan (—) and photograph (—) images compared to our primary results [12] (dotted lines)

In addition to this performance measure, we tried to evaluate the robustness of our approach by comparing the number of leaflets we estimate with the actual number on the leaves. This estimation is performed along the evolution of the compound leaf model and is of crucial importance for the

following of the process. It also has a visual interest for the user witnessing its evolution, and it is beneficial from this point of view that this phase sticks with the actual visible content of the image.

The 1040 images were manually labelled with their number of leaflets, and the estimated number compared with this ground truth. Our method deals with overlapping leaflets, which is basically impossible to do without introducing a kind of model. For the sake of comparison, we also performed a concurrent, non model-based, estimation of this number, by thresholding the color dissimilarity map using Otsu’s method, cleaning the binary image with some mathematical morphology operations, and counting the connected components.

We measured the performance of both approaches by computing the mean squared error (MSE), in a global fashion first, and also for each true number of leaflets. When the thresholding based method has a MSE of 28,67, our estimation reaches 13,11. The error by number of leaflets presented in Figure 5 allows a finer analysis and shows that our approach has a significant advantage when the number of leaflets and the subsequent probability of overlap becomes high.

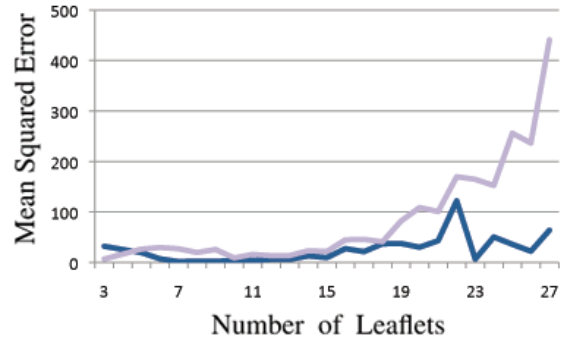


Fig. 5. Mean Squared Error of the leaflet estimation for our method (—) and the Otsu based method (—)

5. CONCLUSIONS & PERSPECTIVES

The approach presented in this paper introduces a model-based method to solve a complex problem of image understanding, providing a high-level interpretation of the image content. The description of compound leaves it produces appears to be more accurate than what can be obtained with less dedicated approaches, through the use of prior botanical knowledge, and constitutes a great improvement compared to our primary results.

The performance in terms of species identification is very satisfying, and the fact that high-level representations are used throughout the process makes an explanatory recognition process using botanical semantics possible. The introduction of this method in the existing mobile application Folia is engaged, and will constitute a nice extension towards a more general tree identification helping tool.

6. REFERENCES

- [1] P. Belhumeur, D. Chen, S. Feiner, D. Jacobs, W. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. Sheorey, S. White, and L. Zhang, "Searching the world's herbaria: A system for visual identification of plant species," in *ECCV*, 2008.
- [2] C.-H. Teng, Y.-T. Kuo, and Y.-S. Chen, "Leaf segmentation, its 3d position estimation and leaf classification from a few images with very close viewpoints," in *ICIAR*, 2009.
- [3] X.F. Wang, D.S. Huang, J.X. Du, X. Huan, and L. Heutte, "Classification of plant leaf images with complicated background," *Applied Mathematics and Computation*, 2008.
- [4] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [5] A. Yuille, P. Hallinan, and D. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 99–111, 1992.
- [6] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active shape models-their training and application," *CVIU*, vol. 61, no. 1, pp. 38–59, 1995.
- [7] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr, "Diffusion snakes: introducing statistical shape knowledge into the mumford-shah functional," *International Journal Of Computer Vision*, vol. 50, pp. 295–313, 2002.
- [8] F. Mokhtarian and S. Abbasi, "Matching shapes with self-intersections: Application to leaf classification," *IEEE Transactions on Image Processing*, vol. 13, no. 5, 2004.
- [9] A. Arora, A. Gupta, N. Bagmar, S. Mishra, and A. Bhat-tacharya, "A plant identification system using shape and morphological features on segmented leaflets," in *CLEF (Notebook Papers/Labs/Workshop)*, 2012.
- [10] C. Caballero and M.C. Aranda, "Plant species identification using leaf image retrieval," in *CIVR*, 2010, pp. 327–334.
- [11] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin, "Guiding active contours for tree leaf segmentation and identification," in *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
- [12] G. Cerutti, V. Antoine, L. Tougne, J. Mille, L. Valet, A. Vacavant, and D. Coquin, "Reves participation - tree species classification using random forests and botanical features," in *CLEF (Notebook Papers/Labs/Workshop)*, 2012.
- [13] G. Cerutti, L. Tougne, A. Vacavant, and D. Coquin, "A parametric active polygon for leaf segmentation and shape estimation," in *ISVC*, 2011.
- [14] T. Brox and J. Weickert, "Level set segmentation with multiple regions," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213–3218, 2006.
- [15] Y. Shi and W.C. Karl, "A real-time algorithm for the approximation of level-set-based curve evolution," *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 645–656, 2008.
- [16] H. Goëau, P. Bonnet, A. Joly, D. Barthelemy, N. Bouje-maa, and J.-F. Molino, "The imageclef 2012 plant image identification task," in *ImageCLEF 2012 Working Notes*, 2012.